Probability: Variance[1]

- **Variance and Standard Deviation.**

  The expectation of a random variable is some sort of an "average behavior" of a random variable. However, the true value of a random variable may be no where close to the expectation. For instance, consider a random variable which takes the value 10000 with probability $1/2$, and -10000 with probability $1/2$. What is $\mathbf{Exp}[X]$? Yes, it is 0. Thus, there is significant *deviation* of X from its expectation.

  The variance and standard deviation try to capture this deviation. In particular, the variance of a random variable is the *expected value of the square of the deviation.*

> Let X be a random variable. The variance of X is defined to be
>
> $$\mathbf{Var}[X] := \mathbf{Exp}\left[(X - \mathbf{Exp}[X])^2\right]$$
>
> That is, if we define another random variable $D := (X - \mathbf{Exp}[X])^2$, then $\mathbf{Var}[X]$ is the expected value of this new deviation random variable D.
>
> The *standard deviation* $\sigma(X)$ is defined to be $\sqrt{\mathbf{Var}(X)}$.

Theorem 1. $\mathbf{Var}[X] = \mathbf{Exp}[X^2] - (\mathbf{Exp}[X])^2$.

Proof.

$$\mathbf{Var}[X] = \mathbf{Exp}[(X - \mathbf{Exp}[X])^2] = \mathbf{Exp}[X^2 - 2X\,\mathbf{Exp}[X] + (\mathbf{Exp}[X])^2]$$

Then, we apply linearity of expectation to get

$$\mathbf{Var}[X] = \mathbf{Exp}[X^2] - 2\,\mathbf{Exp}[X]\cdot\mathbf{Exp}[X] + (\mathbf{Exp}[X])^2 = \mathbf{Exp}[X^2] - (\mathbf{Exp}[X])^2$$

□

A useful corollary (something we observed in the last lecture notes):

Theorem 2. For any random variable $\mathbf{Exp}[X^2] \geq (\mathbf{Exp}[X])^2$.

Proof. $\mathbf{Var}[X]$ is the expected value of $(X - \mathbf{Exp}[X])^2$. That is, $\mathbf{Var}[X]$ is the expected value of a random variable which is always non-negative. In particular, $\mathbf{Var}[X]$ is non-negative. Which in turn means $\mathbf{Exp}[X^2] - (\mathbf{Exp}[X])^2 \geq 0$. Rearranging implies the corollary. □

[1] *Lecture notes by Deeparnab Chakrabarty. Last modified : 28th Aug, 2021*
These have not gone through scrutiny and may contain errors. If you find any, or have any other comments, please email me at deeparnab@dartmouth.edu. Highly appreciated!

Examples

- *Roll of a die.* Let X be the roll of a fair 6-sided die. We know that $\mathbf{Exp}[X] = 3.5$. To calculate the variance, we can use the deviation $D := (X - \mathbf{Exp}[X])^2 = (X - 3.5)^2$. Usinhg this, we get

$$\mathbf{Var}[X] = \mathbf{Exp}[D] = \frac{1}{6}\left((2.5)^2 + (1.5)^2 + (0.5)^2 + (0.5)^2 + (1.5)^2 + (2.5)^2\right) = \frac{35}{12}$$

- *Toss of a biased coin.* Let X be a Bernoulli random variable taking value 1 if a coin tosses heads, and 0 otherwise. Suppose the probability of heads was p. Recall, $\mathbf{Exp}[X] = p$. Also note since X is a indicator random variable, $X^2 = X$. Thus, $\mathbf{Exp}[X^2] = p$ as well. We can calculate the variance as

$$\mathbf{Var}[X] = \mathbf{Exp}[X^2] - (\mathbf{Exp}[X])^2 = p - p^2 = p(1-p)$$

- *Indicator Random Variable.* Using the above toss of a biased coin example, we see that for any event $\mathcal{E}$, the variance of the indicator random variable is

$$\mathbf{Var}[\mathbf{1}_{\mathcal{E}}] = \mathbf{Pr}[\mathcal{E}] \cdot (1 - \mathbf{Pr}[\mathcal{E}]) = \mathbf{Pr}[\mathcal{E}] \cdot \mathbf{Pr}[\neg\mathcal{E}]$$

Theorem 3. If X is a random variable, and c is a "scalar" (a constant), then $Z = c \cdot X$ is another random variable. $\mathbf{Var}[c \cdot X] = c^2 \cdot \mathbf{Var}[X]$.

Proof.

$$\mathbf{Var}[c \cdot X] = \mathbf{Exp}[c^2X^2] - (\mathbf{Exp}[cX])^2 = c^2\,\mathbf{Exp}[X^2] - c^2\,(\mathbf{Exp}[X])^2 = c^{\cdot}\,\mathbf{Var}[X]$$

□

The next theorem is a *linearity of variance* result for ***independent*** random variables.

Theorem 4. For any two ***independent*** random variables X and Y, let $Z := X + Y$. Then,

$$\mathbf{Var}[Z] = \mathbf{Var}[X] + \mathbf{Var}[Y]$$

Proof. By definition of variance, we get

$$\mathbf{Var}[X+Y] = \mathbf{Exp}[(X+Y)^2] - (\mathbf{Exp}[X] + \mathbf{Exp}[Y])^2 \tag{1}$$

Now, we expand the first term in the RHS to get

$$\begin{aligned}
\mathbf{Exp}[(X+Y)^2] &= \mathbf{Exp}[X^2 + 2XY + Y^2] \\
&= \mathbf{Exp}[X^2] + 2\,\mathbf{Exp}[XY] + \mathbf{Exp}[Y^2] && \text{Linearity of Expectation} \\
&= \mathbf{Exp}[X^2] + 2\,\mathbf{Exp}[X]\,\mathbf{Exp}[Y] + \mathbf{Exp}[Y^2] && \text{Since } X \text{ and } Y \text{ are independent.}
\end{aligned} \tag{2}$$

Next, we expand the second term in the RHS of (1), to get

$$(\mathbf{Exp}[X] + \mathbf{Exp}[Y])^2 = (\mathbf{Exp}[X])^2 + 2\,\mathbf{Exp}[X]\,\mathbf{Exp}[Y] + (\mathbf{Exp}[Y])^2 \tag{3}$$

Subtracting (3) from (2), we get

$$\begin{aligned} \mathbf{Var}[X+Y] &= \left(\mathbf{Exp}[X^2] - (\mathbf{Exp}[X])^2\right) + \left(\mathbf{Exp}[Y^2] - (\mathbf{Exp}[Y])^2\right) \\ &= \mathbf{Var}[X] + \mathbf{Var}[Y] \end{aligned} \tag{4}$$

□

We can generalize the above proof to many random variables. In particular, we can say that if $X_1, X_2, \ldots, X_k$ are mutually independent random variables, then the variance of the sum is the sum of the variances. However, we *don't need mutual independence*. Pairwise independence suffices. The proof is given as a solution to the UGP; perhaps you can try it. There is nothing more than the algebra above except there are k things adding up.

Theorem 5. For any k ***pairwise independent*** (and therefore also for mutually independent) random variables $X_1, X_2, \ldots, X_k$,

$$\mathbf{Var}\left[\sum_{i=1}^{k} X_i\right] = \sum_{i=1}^{k} \mathbf{Var}[X_i]$$